

# A SOCIOPHONETIC ANALYSIS OF /l/ DARKNESS AND LEBANESE AUSTRALIAN ETHNIC IDENTITY IN AUSTRALIAN ENGLISH

Josh Clothier

University of Melbourne & ARC Centre of Excellence for the Dynamics of Language  
joshuajc@unimelb.edu.au

## ABSTRACT

Ethno-cultural varieties of Australian English are expected to grow in the 21<sup>st</sup> century [10], yet we know little about the phonetic detail of the linguistic repertoires of speakers from the many ethno-cultural groups in Australia. In Australian English, /l/ has been said to be dark in all word positions; however, this has not been thoroughly tested to date. This paper compares the acoustic properties of /l/ produced by Anglo-Celtic Australians (N = 20), and Lebanese Australians (N = 30), who represent the 9<sup>th</sup> largest ethno-cultural group in Australia.

Results from a wordlist task show that word position has the strongest effect on /l/ darkness, such that /l/ is darker word-finally than word-initially, as seen in other varieties of English. This positional effect is observed for both speaker groups, to differing degrees. Further, for Lebanese Australians variability in /l/ darkness is accounted for by gender, ethnic identity, and social networks.

**Keywords:** Australian English, sociophonetics, ethnic identity, laterals (/l/), acoustic phonetics.

## 1. INTRODUCTION

The Australian English (AusE) accent is generally argued to be relatively homogenous [11]. While it has previously been described as varying along a continuum of broadness, more recently, Cox has argued that we will see less extremes of that continuum, and that there will be an increase in ethnocultural variability in AusE [10].

Australia is home to over 300 distinctive ethnic groups, as avowed in the 2016 census [1]. Among the largest of these groups are people who share Lebanese ethnic heritage. The heritage language of this group, Arabic<sup>1</sup>, is the second most commonly spoken “home language” [1] (third if we include English), and Lebanese Australians are the largest group who claim Arabic as a heritage language (HL) [1].

In Australia, there is a strong background of ethnolinguistic diversity [1], but relative homogeneity [11], particularly as compared to other English-speaking countries such as the UK and the USA [11]. Like varieties within those countries, the alveolar lateral approximant, /l/, in AusE is said to have both light (or clear) and dark allophones that are positionally dis-

tributed [11]; however, while /l/ has also been argued to be dark in all contexts [31] this has not been tested empirically to any great extent.

Previous acoustic phonetic analysis of Lebanese Australians’ speech shows prosodic differences compared to “mainstream” AusE [12], as well as complex sociophonetic variation in VOT which varies as a function of gender, social network, degree of religious affiliation, and ethnic identity [8, 9].

### 1.1. /l/ darkness in English and Arabic

In other varieties of English which exhibit varying degrees of /l/ clearness-darkness, the production of /l/ by minority-ethnic and bilingual speakers has been investigated [e.g., 19, 20]. Of considerable relevance to the present study, Khattab [19] examined the F1 and F2 of bilingual and monolingual Lebanese and Yorkshire English speakers. As Khattab points out, English and Arabic—or, in the present case, Arabic HL speakers—“constitute an interesting pair for comparison”. This is also the case here since AusE is argued to be a particularly dark /l/ variety, and because Arabic /l/ is typically clear in all but a few quite specific contexts [19]. This contrasting set of potential /l/ realisations with conflicting accounts of distribution provides an attractive intersection for analysis.

While previous ethnolinguistic research sought to delineate “ethnolects”, focus has turned to examining speakers’ ethnolinguistic repertoires, acknowledging that speakers who share the same ethnicity might not share the same degree of identification with their ethnicity, or the same kinds of practices—including linguistic practices—related to their ethnicities [3]. Furthermore, a single macro-sociological label denoting the speaker’s ethnicity has become insufficient in capturing the many ways speakers may identify ethnically. Thus, Hall-Lew and Wong [15] argue for categorisations that take more detail of ethnic *identity* into account. Hoffman & Walker [17] provide such a measure, which, they argue, “[combines] subjective/emic and objective/etic approaches to the ethnic categorization of speakers”. The application of this approach in this study is taken up in section 2.2.2.

This paper examines the relationships between /l/ darkness as measured by F2-F1, ethnic heritage, and ethnic identity. The research questions are:

1. Are there positional effects of /l/ darkness in this sample of Australian Englishes?

2. Are there differences in F2-F1 between speakers with Lebanese and Anglo-Celtic ethnic heritage?
3. Within the group of Lebanese Australian speakers, what sociophonetic properties characterise /l/ variation?

## 2. METHOD

### 2.1. Speakers

Data come from two sources: a Lebanese Australian corpus collected by the author, and the AusTalk corpus [13]. Twenty speakers (F = 10, M = 10) were selected from the AusTalk corpus who satisfied the following criteria: raised (and acquired AusE) in the greater Melbourne Metropolitan region; aged 18-30 at the time of recording; and, provided cultural heritage information indicating categorisation as “mainstream” AusE speakers would be appropriate (i.e., heritage is Anglo-Celtic Australian; no information about parents or grandparents from a non-majority ethnicity).

Speakers in the Lebanese Australian corpus (N = 30) are age matched to the young-adult group of the AusTalk corpus (i.e., 18-30y at time of recording) and similarly were raised and acquired AusE in the Melbourne Metropolitan region but have parents and/or grandparents who were born in Lebanon.

### 2.2. Materials

#### 2.2.1. Phonetic data

Data used in this study were elicited using the AusTalk [13] wordlist task, which contains a total of 322 words. Within the wordlist, there are six words with /l/ in absolute word initial position and 23 with /l/ in absolute final position in stressed syllables. Speakers produced three repetitions of each word in the list (although some tokens are excluded due to incidental poor recording quality or speaker production issues). For this analysis, we use all six of the words with initial /l/ and eight of the words with final /l/. Ongoing work will analyse all of the words.

Data for the Lebanese Australian corpus were recorded in a sound treated recording booth at the researcher’s institution using Charter Oak E700 Microphones with Aphex 1100 MKii preamp and a BSS DPR-402 compressor. Data were recorded using a Digidesign 003 rack firewire soundcard using Samplitude Pro X Suite as the software, and were recorded at 16 bit 44.1 KHz. Details of the AusTalk recording protocols are available in [13]. AusTalk data used in this study come from the local subset, which were recorded in the same sound treated recording

booth as the Lebanese Australian corpus by the author in many cases.

From the data selected for this paper, there were 245 vocalized tokens identified auditorily [14]. These were excluded from the analysis, as it was decided that truly vocoid tokens could contaminate the analysis of clear and dark /l/. Thus, we are left presently with 3,969 tokens for analysis (see Table 1. below)

**Table 1:** Distribution of /l/ tokens used in the analysis

Group		N
Anglo-Celtic	Female	741
	Male	657
Lebanese	Female	1262
	Male	1309
	Total	3969

#### 2.2.2. Ethnic identity measure

Lebanese-Australian participants completed a questionnaire designed to quantify ethnic identity, from a constructivist, social psychological perspective. This questionnaire combines items from [17] and [27] and was presented online. Items are primarily Likert-type and enable derivation of an overall score and sub-scores (each out of five) for distinct aspects of ethnic identity. These scores provide means of quantifying the gradient nature of ethnic identity.

### 2.3. Analysis

#### 2.3.1. Segmentation

Data were first automatically segmented at the “phoneme” level using the AusE model in WebMAUS [21]. In order to prepare for later analysis following [5], data were then labelled in Praat [4] for four intervals: lateral onset, lateral steady state, lateral offset, and vowel steady state; however, for the present analysis, we are only interested in the lateral steady state. Following [5], the primary cue for segmentation was F2, as “an approximately steady region of F2 [is] visible on the spectrogram” [17]. Nevertheless, since F1 and F3 also show somewhat consistent patterns in laterals [23, 30], these were used as secondary cues when F2 was insufficient.

After segmentation in Praat [4], data were transferred to emuR [33]. Extracted formant tracks were corrected in the EMU-webApp [32], and F1 and F2 were extracted at the midpoint of the lateral steady state.

While F3 has been shown to be important in the analysis of /l/ darkness [e.g., 28], this will be pursued in future studies.

### 2.3.2. F2-F1 metric

The phenomenon of /l/ darkness lies not just in the values of formants independently, but in the relationships between the formants. F1 is lower and F2 higher for clearer /l/s, whereas F2 is lower (due to dorsal retraction), and F1 is higher in darker /l/s. The F2-F1 metric is used to capture this relationship and is commonly used in studies of /l/ darkness [18, 20, 24], thus also allowing for cross-linguistic and cross-dialectal comparison with reported values.

### 2.3.3. Bark normalization

Formant frequencies (Hz) were Bark normalized in keeping with the common practice in similar studies [e.g., 19, 20]. Bark normalization is based on the principle that, since /l/ darkness is a perceptual phenomenon, a perceptual scale like Bark is more appropriate for providing numerical descriptions of the phenomenon than raw Hz frequencies [c.f., 6].

### 2.3.4. Statistical analysis

Statistical analyses were performed using linear mixed effects regression modelling in **lme4** [2] in **RStudio** [29] with Microsoft R Open [25]. Model selection was aided by the *step* function in the **lmerTest** package [22], with best-fitting models confirmed using likelihood ratio tests. Post-hoc analysis was performed using Bonferroni adjusted *p* values, and only results with *p* values < .05 are treated as significant.

## 3. RESULTS

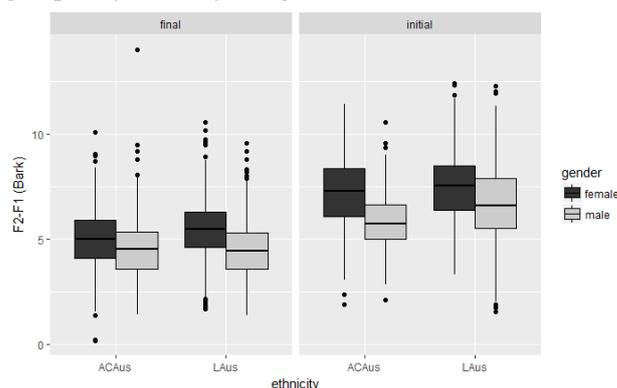
Figure 1 shows boxplots of F2-F1 for /l/ grouped by position, gender, and ethnicity. According to the model of best fit for these data, there are significant main effects of gender, position, and ethnicity, and interactions between gender × ethnicity, gender × position, and ethnicity × position.

As can be seen in figure 1, and confirmed by the modelling, women have higher F2-F1 (Bark), suggestive of clearer /l/, overall compared with men of their same ethnicity, *p* < .001. Likewise, Lebanese Australians have higher F2-F1 (Bark) (clearer /l/) than Anglo-Celtic Australians, *p* = .0429, and initial /l/ have significantly higher F2-F1 (Bark) (are clearer) than final /l/ (which are darker), *p* < .001

The gender × ethnicity interaction confirms the main effect of both gender and ethnicity; Lebanese Australian women have higher F2-F1 (Bark) (clearer /l/) than both Anglo-Celtic Australian men, *p* < .001, and Lebanese Australian men, *p* = .009

The gender × position interaction is significant for all combinations of the terms' levels. This enables the ordering of a cline from lowest to highest F2-F1

**Figure 1:** F2-F1 values (Bark) across word positions grouped by ethnicity and gender.



(Bark) values: (darker /l/) male final <\* male initial <\* female final <\* female initial (clearer /l/).

The ethnicity × position interaction affirms that for both ethnicity groups, initial /l/ have higher F2-F1 (Bark) than do final /l/ (Lebanese Australian, *p* < .001; Anglo-Celtic Australian, *p* < .001).

### 3.1. Results: Lebanese Australians' /l/ variability

The output of the best fitting model to account for sociophonetic variability in Lebanese Australians' /l/ production, as measured by F2-F1 (Bark) is shown in Table 2 below.

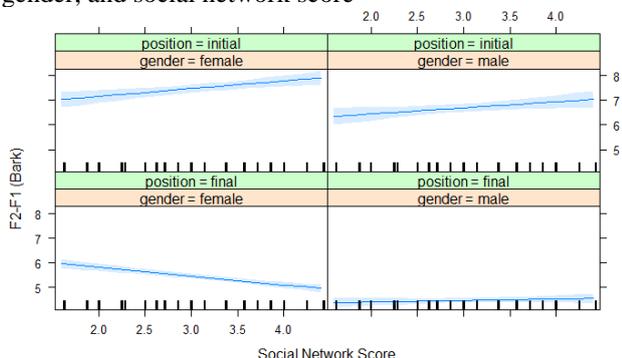
**Table 2:** Final linear mixed effects model for F1-F2 (Bark) in /l/ for both speaker groups. Random intercepts are speaker and word.

	$\beta$	SE	t	<i>p</i>
(Intercept)	6.54	0.64	10.22	<.001
Gender: male	-2.25	0.95	-2.36	.026
network subscore	-0.36	0.19	-1.89	.069
position: initial	0.00	0.40	0.00	.999
gender: male × network subscore	0.42	0.30	1.94	.174
gender: male × position: initial	1.66	0.42	3.95	<.001
network subscore × position: initial	0.67	0.08	7.95	<.001
gender: male × network subscore × position: initial	-0.48	0.13	-3.65	<.001

The three-way interaction identified by this model is also shown in figure 2. Here we see that for initial /l/ produced by Lebanese Australian women and men, as the density of their social network increases, so too does their F2-F1 (Bark) (i.e., the /l/s become clearer. In final position, there is no effect for men, but for women, the opposite effect occurs; as

their social network becomes more Lebanese, their final /l/s become darker.

**Figure 2:** LMER model fits for F2-F1 (Bark) values in Lebanese Australian /l/: interactions between position, gender, and social network score



#### 4. DISCUSSION

The F2-F1 analysis presented here shows a consistent effect of position in the word on F2-F1. This suggests that there is a positional effect for clear-dark /l/ in AusE, whereby initial /l/s are clearer and final /l/s are darker. This effect is robust: it is found across two contemporary speaker groups and, while there are complexities to the data (discussed further below), the effect of position remains significant in each subsequent analysis and interaction.

Between the two groups, there is no major difference in the production of /l/ (based on F2-F1 (Bark)): it is only when we decompose the samples and draw comparisons on the basis of position and gender that differences emerge, and, here, we find that, Lebanese Australian women tend to have higher F2-F1 (Bark) values, suggesting clearer /l/ productions. However, in terms of position, both the Lebanese Australian and Anglo-Celtic Australian groups have significantly clearer initial /l/s and significantly darker final /l/s.

Variability among Lebanese Australians' /l/s is explained by a three-way interaction between gender, social network score, and word position. This accords with previous findings for Lebanese Australians, where male speakers' voiceless stop VOT decreased as social network score increases. Further, social networks are known to be an important factor in sociolinguistic variation [26], and play a key role in the development of ethnolinguistic repertoires [7].

Considering the pattern shown by the Lebanese Australian women, in both the initial and final positions, this pattern is particularly interesting if we consider their VOT results [8, 9], as it suggests the emergence of a larger pattern of sociophonetic behaviour. For word initial /l/, as the social network becomes more densely Lebanese, /l/ becomes clearer. For word final /l/, the reverse is true: the increase in Lebanese social network density is associated with *darker* /l/.

Lebanese Australian women with strong, dense ties to the Lebanese community make a sharp distinction between dark /l/ in final position and clear /l/ in initial position, showing a “classic” English pattern, and certainly no evidence of “substratum transfer” for this variable. A similar effect was found in their VOT system: for voiceless /p t k/, as their sense of ethnic identity increased, so too did their positive VOT [8]. This serves as a crucial reminder in the study of ethnicity: the linguistic behaviour of men and women must be considered separately—not simply because of differences in vocal tract lengths, but because men and women are socialised into their ethnicities differently, and, likewise, the linguistic performance of their ethnicities will be different [16].

#### 5. CONCLUSION

This sociophonetic study of /l/ in Australian Englishes as spoken by those with mainstream, Anglo-Celtic, and those with Lebanese ethnic identities shows that, in accordance with [11] AusE does have positionally distributed clear and dark “variants” of /l/. Future comparison with published values on other varieties will allow greater clarity on AusE’s status as a “dark” /l/ variety as argued by [31].

We find no overall group differences in /l/ darkness between Lebanese and Anglo-Celtic Australians; however, Lebanese Australian women have clearer /l/s than other groups. Considering this alongside previous gender-based findings will be enlightening, as will further exploration of individual differences in speakers from both ethnic groups.

Future work, as part of the larger project, will involve analysis of dynamics of formant trajectories in VL and LV sequences, and analysis of the vowel system, to enhance the present analysis.

The results of this study raise additional questions for future research. Studying the social perception of this variable would be enlightening: what does a Lebanese Australian /l/ sound like to speakers within and outside of the community? Are results reported here socially meaningful for speakers? This would provide a more definitive set of answers about the sociophonetic status of /l/ in (Lebanese) Australian English.

#### 6. ACKNOWLEDGEMENTS

This research is supported by an Australian Government Research Training Program scholarship. I also thank my PhD supervisors, Janet Fletcher, Debbie Loakes, and Barb Kelly, and the speakers who so generously gave of their time to enable this work.

## 7. REFERENCES

- [1] Australian Bureau of Statistics, 2016. *Cultural Diversity (2016)*, Table Builder. Findings based on use of ABS TableBuilder data.
- [2] Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4, *J. Statistical Software*, 67, 1-48.
- [3] Benor, S. B., 2010. Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity<sup>1</sup>, *J. Sociolinguistics*, 14, 159-183.
- [4] Boersma, P., Weenink, D., 2017. Praat: doing phonetics by computer. 6.0.29 ed.
- [5] Carter, P., Local, J., 2007. F2 variation in Newcastle and Leeds English liquid systems, *JIPA*, 37, 183.
- [6] Carter, P. G., Structured variation in British English liquids: the role of resonance, PhD Unpublished Thesis, Department of Language and Linguistic Science, University of York, York, UK, 2002.
- [7] Cheshire, J., Kerswill, P., Fox, S., Torgersen, E., 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English, *J. Sociolinguistics*, 15, 151-196.
- [8] Clothier, J., Fletcher, J., Loakes, D., 2018. The relationship between English voice onset time and ethnic identity among Lebanese-Australians in Melbourne, Paper presented at the Sociolinguistics Symposium 22, Auckland, NZ
- [9] Clothier, J., Loakes, D. 2018. Coronal stop VOT in Australian English: Lebanese Australians and mainstream Australian English, in *Proc 17th SST*, J. Epps, J. Wolfe, J. Smith, and C. Jones, Eds. Sydney, Australia: UNSW.
- [10] Cox, F., 2006. Australian English pronunciation into the 21st century, *Prospect*, 21, 3-21.
- [11] Cox, F., Fletcher, J., 2017. *Australian English Pronunciation and Transcription*. Cambridge University Press.
- [12] Cox, F., Palethorpe, S., 2011. Timing differences in the VC rhyme of standard Australian English and Lebanese Australian English, *Proc. 17th ICPHS Hong Kong*, 528-531.
- [13] Estival, D., Cassidy, S., Cox, F., Burnham, D., 2014. AusTalk: an audio-visual corpus of Australian English, presented at the 9th Language Resources and Evaluation Conference, Reykjavik, Iceland.
- [14] Hall-Lew, L., Fix, S., 2012. Perceptual coding reliability of (L)-vocalization in casual speech data, *Lingua*, 122, 794-809.
- [15] Hall-Lew, L., Wong, A. W.-m., 2014. Coding for Demographic Categories in the Creation of Legacy Corpora: Asian American Ethnic Identities, *Language and Linguistics Compass*, 8, 564-576.
- [16] Hall, K., Bucholtz, M. 1995. From Mulatta to Mestiza: Language and the Reshaping of Ethnic Identity, in *Gender articulated: Language and the socially constructed self*, K. Hall and M. Bucholtz, Eds.: Routledge, 351-374.
- [17] Hoffman, M. F., Walker, J. A., 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English, *LVC*, 22, 37-67.
- [18] Holmes-Elliott, S., Smith, J., 2018. Dressing down up north: DRESS-lowering and /l/ allophony in a Scottish dialect, *LVC*, 30, 23-50.
- [19] Khattab, G. 2011. Acquisition of Lebanese Arabic and Yorkshire English /l/ by bilingual and monolingual children, in *Instrumental Studies in Arabic Phonetics*, Z. Majeed Hassan and B. Heselwood, Eds.: Benjamins.
- [20] Kirkham, S., 2017. Ethnicity and phonetic variation in Sheffield English liquids, *JIPA*, 47, 17-35.
- [21] Kisler, T., Reichel, U., Schiel, F., 2017. Multilingual processing of speech via web services, *Computer Speech & Language*, 45, 326-347.
- [22] Kuznetsova, A., Bruun Brockhoff, P., Haubo Bojesen Christensen, R., 2017. Package: Tests in Linear Mixed Effects Models, *J. Statistical Software*, 82, 1-26.
- [23] Ladefoged, P., Maddieson, I., 1996. *The Sounds of the World's Languages*. Wiley.
- [24] Mackenzie, S., De Decker, P., Pierson, R., Year. An Acoustic and Articulatory Study of /l/ Allophony in Newfoundland English, in *Proc. 18th ICPHS Glasgow*.
- [25] Microsoft, R Core Team, 2018. Microsoft R Open. 3.5.0 ed.
- [26] Milroy, L., Llamas, C. 2013. Social networks, in *The handbook of language variation and change*, J. K. Chambers and N. Schilling, Eds. West Sussex, UK: Wiley-Blackwell, 407-427.
- [27] Phinney, J. S., Ong, A. D., 2007. Conceptualization and measurement of ethnic identity: Current status and future directions, *J. Counseling Psychology*, 54, 271-281.
- [28] Popescu, A., 2016. F3 variability in allophones of /l/: Acoustic-articulatory relations, *J. Acoust. Soc. Am.*, 140, 3222-3222.
- [29] RStudio Team, 2018. RStudio: Integrated Development for R. RStudio, Inc. vol. desktop, 1.1.463 ed. Boston, MA.
- [30] Stuart-Smith, J. *et al.*, 2015. A Dynamic Acoustic View of Real-Time Change in Word-Final Liquids in Spontaneous Glaswegian, *Proc. 18th ICPHS Glasgow*.
- [31] Wells, J., 1982. *Accents of English: Beyond the British Isles*. Cambridge University Press.
- [32] Winkelmann, R., Harrington, J., Jansch, K., 2017. EMU-SDMS: Advanced speech database management and analysis in R, *Computer Speech & Language*, 45, 392-410.
- [33] Winkelmann, R., Jaensch, K., Cassidy, S., Harrington, J., 2018. emuR: Main Package of the EMU Speech Database Management System. R package version 1.0.0 ed.

<sup>1</sup> While Lebanese Arabic or, simply, Lebanese, is a more accurate descriptor, this is not adequately accounted for in the Census.